

## Abstract

We determine the range of Furstenberg entropy for stationary ergodic actions of nonabelian free groups by an explicit construction involving random walks on random coset spaces.

# Random walks on random coset spaces with applications to Furstenberg entropy

Lewis Bowen\*  
Texas A&M University

April 24, 2012

**Keywords:** Furstenberg entropy,  $\mu$ -entropy, random walk entropy, stationary dynamical systems, Poisson boundaries

**MSC:**37A40, 37A15, 37A50, 60G10

## Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>Poisson boundaries of random walks on coset spaces</b>	<b>5</b>
2.1	Maps . . . . .	6
2.2	The space of subgroups . . . . .	6
2.3	Actions . . . . .	7
<b>3</b>	<b>Entropy formulae</b>	<b>7</b>
<b>4</b>	<b>Results for the free group</b>	<b>12</b>
4.1	A continuity criterion . . . . .	13
4.2	A covering space construction . . . . .	16
4.3	Paths of IRS's . . . . .	18
4.4	Entropies of boundaries of quotient groups . . . . .	21

---

\*supported in part by NSF grant DMS-0968762, NSF CAREER Award DMS-0954606 and BSF grant 2008274

# 1 Introduction

Let  $\mu$  be a Borel probability measure on a locally compact group  $G$ . An action of  $G$  on a probability space  $(X, \eta)$  is  $\mu$ -stationary if  $\eta = \mu * \eta$  where

$$\mu * \eta := \int g_* \eta \, d\mu(g)$$

is the *convolution* of  $\mu$  with  $\eta$ . There is significant interest in understanding the structure of stationary actions and their connections with random walks [Fu63b, Fu71, Fu72, Fu80], rigidity theory [NZ99, NZ00, NZ02a, NZ02b, Ne03] and classification of invariant measures [BFLM11, BQ09, BQ11a, BQ11b]. A general structure theory is presented in [FG10].

Stationary systems are abundant; indeed every continuous action of  $G$  on a compact metric space admits a stationary measure. However tractable examples, other than Poisson boundaries and measure-preserving actions, are somewhat lacking. One of the main contributions of this paper is the construction of new examples.

The *Furstenberg entropy* or  $\mu$ -entropy of a  $\mu$ -stationary action of  $G$  on a probability space  $(X, \eta)$  is a fundamental invariant defined in [Fu63a] by

$$h_\mu(X, \eta) := \iint -\log \frac{d\eta \circ g}{d\eta}(x) \, d\eta(x) \, d\mu(g).$$

By Jensen's inequality this entropy is always nonnegative. It equals zero if and only if the action is measure-preserving. One of the main results of [NZ00] and [NZ02a] is that if  $G$  is a connected higher rank real semisimple Lie group with finite center and the action satisfies a certain mixing hypothesis, then this entropy can take on only a finite number of values corresponding with the actions of  $G$  on homogeneous spaces  $(G/Q, \nu_Q)$  where  $Q < G$  is a parabolic subgroup. Indeed, it is shown that any such  $(G, \mu)$ -space is a relatively measure-preserving extension of one of these actions. This is a crucial step in Nevo-Zimmer's proof of the generalized intermediate factor theorem, which constitutes a major generalization of Margulis' normal subgroup theorem.

These results motivate the

**Furstenberg entropy realization problem:** *Given  $(G, \mu)$  what are all possible values of the  $\mu$ -entropy  $h_\mu(X, \eta)$  as  $(X, \eta)$  varies over all ergodic  $\mu$ -stationary actions of  $G$ ?*

In [NZ00], page 323, the authors remark that they do not know the full set of possible values of the Furstenberg entropy for a given  $(G, \mu)$  or even whether this set of values contains an interval (for any non-amenable group  $G$ ). However, they prove that if  $G$  is  $PSL_2(\mathbb{R})$  or a semisimple group of real rank  $\geq 2$  containing a parabolic subgroup that maps onto  $PSL_2(\mathbb{R})$  then infinitely many different values are achieved [NZ00, Theorem 3.4]. It is also proven that if  $G$  has Property (T) then there is an open interval  $(0, \epsilon(\mu))$  containing no values of  $h_\mu(X, \eta)$  for any ergodic  $\mu$ -stationary  $G$ -systems  $(X, \eta)$  [Ne03]. Our main theorem is:

**Theorem 1.1.** *Let  $G = \langle s_1, \dots, s_r \rangle$  be a free group of rank  $2 \leq r < \infty$ ,  $\mu$  be the uniform probability measure on  $\{s_1, \dots, s_r, s_1^{-1}, \dots, s_r^{-1}\}$  and  $h_{\max}(\mu)$  denote the maximum value of the  $\mu$ -entropy over all  $\mu$ -stationary  $G$ -actions  $(X, \eta)$ . Then for every  $t \in [0, h_{\max}(\mu)]$  there exists an ergodic  $\mu$ -stationary  $G$ -action on a probability space  $(X, \eta)$  with  $h_\mu(X, \eta) = t$ .*

To sketch the proof and explain further results, let us recall the notion of Poisson boundary. So consider a locally compact group  $G$  with a probability measure  $\mu$  on  $G$ . Let  $X_1, X_2, \dots$  be a sequence of independent random variables each with law  $\mu$ . The sequence  $\{Z_n\}_{n=1}^\infty$  where  $Z_n := X_1 \cdots X_n$  is the random walk induced by  $\mu$ . The *Poisson boundary* of this random walk, denoted  $(B, \nu)$ , is the space of ergodic components of the time shift on  $(G^\mathbb{N}, \mathbb{P})$  where  $\mathbb{P}$  is the law of the random walk  $\{Z_n\}_{n=1}^\infty$ . Because the time shift commutes with the left-action of  $G$  on  $G^\mathbb{N}$ ,  $G$  acts on the Poisson boundary. This action is  $\mu$ -stationary. It is well-known that  $h_\mu(B, \nu) = h_{\max}(\mu)$  (see e.g. [KV83, §3.2, Corollary 3]).

If  $K < G$  is a closed subgroup, then we may consider the random walk  $\{KZ_n\}_{n=1}^\infty$  on the coset space  $K \backslash G$ . The Poisson boundary of this random walk is the space  $(B_K, \nu_K)$  of ergodic components of the time shift on  $((K \backslash G)^\mathbb{N}, \mathbb{P}_K)$  where  $\mathbb{P}_K$  is the law of the random walk  $\{KZ_n\}_{n=1}^\infty$ . If  $K$  is normal in  $G$ , then  $G$  acts on the left on  $(K \backslash G)^\mathbb{N}$  and this action descends to an action on  $B_K$ . Moreover  $\nu_K$  is  $\mu$ -stationary. Our second main result is:

**Theorem 1.2.** *Let  $(G, \mu)$  be as in Theorem 1.1. Then the set of numbers  $\{h_\mu(B_N, \nu_N) : N < G\}$  is dense in  $[0, h_{\max}(\mu)]$ .*

If  $K$  is not normal in  $G$  then there is no canonical action of  $G$  on  $B_K$ . To remedy this, consider the space  $\text{Sub}_G$  of all closed subgroups of  $G$ .  $G$  acts on this space by conjugation. Let  $\mathcal{M}(\text{Sub}_G)$  denote the space of conjugation-invariant Borel probability measures on  $\text{Sub}_G$ . A random subgroup with law  $\lambda \in \mathcal{M}(\text{Sub}_G)$  is called an *invariant random subgroup* or IRS for short. There has been a recent increase in studies of the action of  $G$  on  $\text{Sub}_G$  and its invariant measures [Bo12, AGV12, Vo12, ABBGNRS11, Ve11, Sa11, Gr11, Ve10, BS06, DS02, GS99, SZ94].

For  $\lambda \in \mathcal{M}(\text{Sub}_G)$ , we consider the random walk  $\{KZ_n\}_{n=1}^\infty$  on the coset space  $K \backslash G$  where  $K < G$  is random with law  $\lambda$  (and  $Z_n$  are as above). The Poisson boundary of this random walk is the space  $(B(\text{Sub}_G), \nu_\lambda)$  of ergodic components of the time shift on  $(\widetilde{\text{Sub}_G}, \mathbb{P}_\lambda)$  where  $\widetilde{\text{Sub}_G}$  is the set of all  $(K; Kg_0, Kg_1, \dots)$  with  $K \in \text{Sub}_G$ ,  $g_0, g_1, \dots \in G$  and  $\mathbb{P}_\lambda$  is the law of  $(K; Kg_0, Kg_1, \dots)$ . The group  $G$  naturally acts on this space and  $\nu_\lambda$  is stationary and ergodic if  $\lambda$  is ergodic.

Incidentally, we will prove a few fundamental results about these random walks in the case of an arbitrary countable discrete group  $G$ . For example, the *random walk entropy* of the walk  $\{Z_n\}_{n=1}^\infty$  on  $G$  is defined to be  $\lim_{n \rightarrow \infty} n^{-1} H(\mu^n)$  where  $\mu^n$  is the  $n$ -fold convolution power of  $\mu$  and  $H(\mu^n) = -\sum_{g \in G} \mu^n(\{g\}) \log \mu^n(\{g\})$ . In [KV83], Kaimanovich and Vershik proved that the random walk entropy equals the Furstenberg entropy of the associated Poisson boundary. In §3 this result is generalized to random walks on the coset space of an invariant random subgroup.

The map which takes  $\lambda \in \mathcal{M}(\text{Sub}_G)$  to  $h_\mu(B(\text{Sub}_G), \nu_\lambda)$  is not continuous in general. For example, consider a decreasing sequence  $\{N_i\}_{i=1}^\infty$  of finite-index normal subgroups with

trivial intersection  $\cap_{i=1}^{\infty} N_i = \{e\}$ . If  $\delta_i \in \mathcal{M}(\text{Sub}_G)$  is the Dirac measure concentrated on  $N_i$  then  $(B(\text{Sub}_G), \nu_{\delta_i}) = (B_{N_i}, \nu_{N_i})$ . Because  $N_i$  has finite index,  $h_{\mu}(B_{N_i}, \nu_{N_i}) = 0$ . However,  $\delta_i$  converges as  $i \rightarrow \infty$  to  $\delta_e$ , the Dirac measure concentrated on the trivial subgroup. Because  $h_{\mu}(B_e, \nu_e) = h_{\max}(\mu) > 0$ , this map is discontinuous. In spite of this discontinuity, we will show that when  $G$  is a free group, there exist paths in  $\mathcal{M}(\text{Sub}_G)$  on which entropy varies continuously and use these paths to establish Theorem 1.1.

**Acknowledgements:** I'd like to thank Amos Nevo for asking me whether Theorem 1.1 is true and for several motivating discussions.

## 2 Poisson boundaries of random walks on coset spaces

Let  $G$  be a locally compact group with a probability measure  $\mu$ . We assume  $\mu$  is *admissible*: its support generates  $G$  as a semigroup and some convolution power  $\mu^n$  is absolutely continuous with respect to Haar measure on  $G$ . The purpose of this section is to set notation and define the Poisson boundary of the  $\mu$ -induced random walk on a coset space  $K \backslash G$ .

Let  $\mathbb{N} := \{1, 2, \dots\}$  and  $\mathbb{N}+1 := \{0, 1, 2, \dots\}$  and  $m : G^{\mathbb{N}+1} \rightarrow G^{\mathbb{N}+1}$  be the multiplication map

$$m(g_0, g_1, g_2, \dots) := (g_0, g_0 g_1, g_0 g_1 g_2, \dots).$$

$\mathbb{P}_g$  is the measure on  $G^{\mathbb{N}+1}$  defined by  $\mathbb{P}_g := m_*(\delta_g \times \mu^{\mathbb{N}})$  where  $\delta_g$  is the Dirac probability measure concentrated on  $\{g\} \subset G$ . We write  $\mathbb{P}$  to denote  $\mathbb{P}_e$  where  $e$  is the identity element.

Let  $K < G$  be a closed subgroup and  $\pi_K : G^{\mathbb{N}+1} \rightarrow (K \backslash G)^{\mathbb{N}+1}$  the quotient map

$$\pi_K(g_0, g_1, g_2, \dots) := (Kg_0, Kg_1, Kg_2, \dots).$$

$\mathbb{P}_{Kg} := (\pi_K)_* \mathbb{P}_g$  denotes the pushforward measure. Of course,  $\mathbb{P}_K := \mathbb{P}_{Ke}$ .

Let  $\sigma : G^{\mathbb{N}+1} \rightarrow G^{\mathbb{N}+1}$  be the shift map:

$$\sigma(g_0, g_1, g_2, \dots) := (g_1, g_2, \dots).$$

Denote the sigma-algebra of shift-invariant Borel subsets of  $G^{\mathbb{N}+1}$  by  $\mathcal{B}(\sigma)$  and let  $B_e$  denote the Borel space  $(G^{\mathbb{N}+1}, \mathcal{B}(\sigma))$ . Let  $\mathbf{bnd} : G^{\mathbb{N}+1} \rightarrow B_e$  be the factor map and  $\nu_g := \mathbf{bnd}_* \mathbb{P}_g$  the pushforward measure on  $B_e$  (for any  $g \in G$ ). The probability space  $(B_e, \nu_e)$  is the *Poisson boundary* of  $(G, \mu)$ .

Similarly, let  $\sigma_K : (K \backslash G)^{\mathbb{N}+1} \rightarrow (K \backslash G)^{\mathbb{N}+1}$  be the shift map:

$$\sigma_K(Kg_0, Kg_1, Kg_2, \dots) := (Kg_1, Kg_2, \dots).$$

Denote the sigma-algebra of shift-invariant Borel subsets of  $(K \backslash G)^{\mathbb{N}+1}$  by  $\mathcal{B}(\sigma_K)$ . Let  $B_K$  be the Borel space  $((K \backslash G)^{\mathbb{N}+1}, \mathcal{B}(\sigma_K))$ ,  $\mathbf{bnd}_K : (K \backslash G)^{\mathbb{N}+1} \rightarrow B_K$  be the factor map and  $\nu_{Kg} := (\mathbf{bnd}_K)_* \mathbb{P}_{Kg}$  the pushforward measure on  $B_K$  (for any  $g \in G$ ). Then  $(B_K, \nu_K)$  is the *Poisson boundary* of  $K \backslash G$  generated by  $\mu$ .

The commutative diagram:

$$\begin{array}{ccc} (G^{\mathbb{N}+1}, \mathbb{P}) & \xrightarrow{\mathbf{bnd}} & (B_e, \nu_e) \\ \downarrow \pi_K & & \downarrow \pi_K \\ ((K \setminus G)^{\mathbb{N}+1}, \mathbb{P}_K) & \xrightarrow{\mathbf{bnd}} & (B_K, \nu_K) \end{array}$$

uses an abuse of notation: we let  $\pi_K$  denote the map from  $G^{\mathbb{N}+1}$  to  $(K \setminus G)^{\mathbb{N}+1}$  as well as the induced map from  $B_e$  to  $B_K$ . Also we let  $\mathbf{bnd}$  denote the map from  $G^{\mathbb{N}+1}$  to  $B_e$  as well as the map from  $(K \setminus G)^{\mathbb{N}+1}$  to  $B_K$  when no confusion can arise.

## 2.1 Maps

For  $(g_0, g_1, g_2, \dots) \in G^{\mathbb{N}+1}$  and  $\gamma \in G$  we let

$$\gamma \cdot (g_0, g_1, g_2, \dots) := (\gamma g_0, \gamma g_1, \gamma g_2, \dots).$$

Because this action commutes with the shift action, it descends to an action of  $G$  on  $B_e$ .

Similarly, for  $(Kg_0, Kg_1, Kg_2, \dots) \in (K \setminus G)^{\mathbb{N}+1}$  and  $\gamma \in G$  we let  $K^\gamma = \gamma K \gamma^{-1}$  and

$$\gamma \cdot (Kg_0, Kg_1, Kg_2, \dots) := (\gamma Kg_0, \gamma Kg_1, \gamma Kg_2, \dots) \in (K^\gamma \setminus G)^{\mathbb{N}+1}.$$

This is not an action on  $(K \setminus G)^{\mathbb{N}+1}$ ; instead it is a map from  $(K \setminus G)^{\mathbb{N}+1}$  to  $(K^\gamma \setminus G)^{\mathbb{N}+1}$ . However it does commute with the shift so each  $\gamma \in \Gamma$  gives rise to a map  $B_K$  to  $B_{K^\gamma}$ .

## 2.2 The space of subgroups

The group  $G$  acts on the set of its closed subgroups  $\text{Sub}_G$  by conjugation. The set  $\text{Sub}_G$  with the topology of uniform convergence on compact subsets is a compact metrizable space. Let  $\mathcal{M}(\text{Sub}_G)$  be the space of all conjugation-invariant Borel probability measures on  $\text{Sub}_G$ .

Let  $\widetilde{\text{Sub}_G} = \{(K; Kg_0, Kg_1, Kg_2, \dots) : K \in \text{Sub}_G, g_0, g_1, g_2, \dots \in G\}$ . Given an invariant measure  $\lambda \in \mathcal{M}(\text{Sub}_G)$ , let  $\mathbb{P}_\lambda$  be the measure on  $\widetilde{\text{Sub}_G}$  whose fiber over  $K \in \text{Sub}_G$  is  $\mathbb{P}_K$ :

$$d\mathbb{P}_\lambda(K; Kg_0, Kg_1, Kg_2, \dots) = d\mathbb{P}_K(Kg_0, Kg_1, \dots) d\lambda(K).$$

Let  $B(\text{Sub}_G) = \{(K; \xi) : K \in \text{Sub}_G, \xi \in B_K\}$ . Let  $\nu_\lambda$  be the measure on  $B(\text{Sub}_G)$  whose fiber over  $K \in \text{Sub}_G$  is  $\nu_K$ :

$$d\nu_\lambda(K; \xi) = d\nu_K(\xi) d\lambda(K).$$

We have the following commutative diagram:

$$\begin{array}{ccc} (\text{Sub}_G \times G^{\mathbb{N}+1}, \lambda \times \mathbb{P}) & \xrightarrow{\mathbf{bnd}} & (\text{Sub}_G \times B_e, \lambda \times \nu_e) \\ \downarrow \pi & & \downarrow \pi \\ (\widetilde{\text{Sub}_G}, \mathbb{P}_\lambda) & \xrightarrow{\mathbf{bnd}} & (B(\text{Sub}_G), \nu_\lambda) \end{array}$$

By abuse of notation we let **bnd** denote both the map from  $\text{Sub}_G \times G^{\mathbb{N}+1}$  to  $\text{Sub}_G \times B_e$  which takes  $(K; g_0, g_1, \dots)$  to  $(K, \mathbf{bnd}(g_0, g_1, \dots))$  as well as the map from  $\widetilde{\text{Sub}}_G$  to  $B(\text{Sub}_G)$  which takes  $(K; Kg_0, Kg_1, \dots)$  to  $(K; \mathbf{bnd}_K(Kg_0, Kg_1, \dots))$ . We also let  $\pi$  denote both the map from  $\text{Sub}_G \times G^{\mathbb{N}+1}$  to  $\widetilde{\text{Sub}}_G$  which takes  $(K; g_0, g_1, \dots)$  to  $(K; Kg_0, Kg_1, \dots)$  as well as the map from  $\text{Sub}_G \times B_e$  to  $B(\text{Sub}_G)$  which takes  $(K, \xi)$  to  $(K, \pi_K(\xi))$ .

## 2.3 Actions

The group  $G$  acts on  $\widetilde{\text{Sub}}_G$  by

$$\gamma(K; Kg_0, Kg_1, \dots) := (K^\gamma; \gamma Kg_0, \gamma Kg_1, \dots) \quad \forall \gamma, g_0, g_1, \dots \in G, K \in \text{Sub}_G.$$

This action commutes with the shift action  $\tilde{\sigma} : \widetilde{\text{Sub}}_G \rightarrow \widetilde{\text{Sub}}_G$  which is defined by

$$\tilde{\sigma}(K; Kg_0, Kg_1, \dots) := (K; Kg_1, \dots).$$

The map  $\mathbf{bnd} : \widetilde{\text{Sub}}_G \rightarrow B(\text{Sub}_G)$  is the quotient of  $\widetilde{\text{Sub}}_G$  with respect to the  $\tilde{\sigma}$ -invariant sigma-algebra. Because the  $G$ -action commutes with this shift it descends to an action on  $B(\text{Sub}_G)$  which is:  $\gamma(K, \xi) = (K^\gamma, \gamma \cdot \xi)$  where  $\gamma \cdot \xi$  is as defined in §2.1. Note that the maps in the diagram above are equivariant with respect to the various  $G$ -actions where the  $G$ -actions on  $\text{Sub}_G \times G^{\mathbb{N}+1}$  and  $\text{Sub}_G \times B_e$  are the diagonal actions.

**Lemma 2.1.** *If  $\lambda \in \mathcal{M}(\text{Sub}_G)$  is ergodic for the  $G$ -action then  $\nu_\lambda$  is also ergodic for the  $G$ -action. Moreover  $\nu_\lambda$  is  $\mu$ -stationary.*

*Proof.* From the diagram above, it follows that  $(\text{Sub}_G \times B_e, \lambda \times \nu_e)$  factors onto  $(B(\text{Sub}_G), \nu_\lambda)$ . Because the Poisson boundary  $(B_e, \nu_e)$  is weakly mixing [AL05] and  $\lambda$  is ergodic,  $(\text{Sub}_G \times B_e, \lambda \times \nu_e)$  is ergodic. Since  $(B(\text{Sub}_G), \nu_\lambda)$  is a factor of an ergodic system, it is also ergodic. The measure  $\lambda \times \nu_e$  is stationary since  $\lambda$  is invariant and  $\nu_e$  is stationary. Since  $(B(\text{Sub}_G), \nu_\lambda)$  is a factor of a stationary system, it is also stationary.  $\square$

## 3 Entropy formulae

In this section, we require  $G$  to be a countable discrete group with an admissible measure  $\mu$ . Our goal in the section is to provide a formula for the  $\mu$ -entropy in terms of the so-called random walk entropy. To explain, we need a few definitions.

We let  $\mu^n$  be the  $n$ -fold convolution of  $\mu$ . In other words, if  $m_n : G^n \rightarrow G$  denotes the multiplication map

$$m_n(g_1, g_2, \dots, g_n) = g_1 g_2 \cdots g_n$$

and  $(G^n, (\times \mu)^n)$  denotes the direct product of  $n$  copies of  $(G, \mu)$  then  $\mu^n = (m_n)_*(\times \mu)^n$ .

For  $K \in \text{Sub}_G$ , let  $\mu_K^n$  be the measure on  $K \backslash G$  given by  $\mu_K^n := (\pi_K)_* \mu^n$  where  $\pi_K : G \rightarrow (K \backslash G)$  is the quotient map. Similarly, if  $g, h \in G$  then  $\mu_{gKh}^n$  is the measure on  $gKg^{-1} \backslash G$  given by

$$\mu_{gKh}^n(E) = \mu^n(\{\gamma \in G : gKh\gamma \in E\}) \quad \forall E \subset gKg^{-1} \backslash G.$$

In general, if  $\omega$  is a probability measure on a finite or countable set  $W$  then the *entropy* of  $\omega$  is

$$H(\omega) := - \sum_{w \in W} \omega(\{w\}) \log(\omega(\{w\}))$$

where by convention  $0 \log(0) = 0$ .

The sequence  $\{H(\mu^n)\}_{n=1}^\infty$  can be shown to be sub-additive. Therefore the limit of  $\frac{H(\mu^n)}{n}$  as  $n \rightarrow \infty$  exists. This limit is called the *random walk entropy* of  $(G, \mu)$ . In ([KV83], Theorem 3.1), it is shown that this coincides with the  $\mu$ -entropy of the Poisson boundary  $(B_e, \nu_e)$ . Analogously, the main result of this section is:

**Theorem 3.1.** *For any invariant measure  $\lambda \in \mathcal{M}(\text{Sub}_G)$ ,*

$$\begin{aligned} h_\mu(B(\text{Sub}_G), \nu_\lambda) &= \lim_{n \rightarrow \infty} \frac{1}{n} \int H(\mu_K^n) d\lambda(K) = \inf_n \frac{1}{n} \int H(\mu_K^n) d\lambda(K) \\ &= \lim_{n \rightarrow \infty} \int H(\mu_K^n) - H(\mu_K^{n-1}) d\lambda(K) = \inf_{n \rightarrow \infty} \int H(\mu_K^n) - H(\mu_K^{n-1}) d\lambda(K). \end{aligned}$$

For  $y \in G^{\mathbb{N}+1}$  or  $y \in (K \setminus G)^{\mathbb{N}+1}$  we let  $y_n$  be the  $n$ -coordinate of  $y$ . So  $y = (y_0, y_1, \dots)$ . We let  $\alpha_n$  be the partition of  $G^{\mathbb{N}+1}$  determined by the condition that  $y, y'$  are in the same partition element if and only if  $y'_i = y_i$  for  $0 \leq i \leq n$ . We let  $\eta_n$  be the partition of  $G^{\mathbb{N}+1}$  determined by the condition that  $y, y'$  are in the same partition element if and only if  $y'_i = y_i$  for  $i \geq n$ . We let  $\tau_n$  be the partition of  $G^{\mathbb{N}+1}$  determined by the condition that  $y, y'$  are in the same partition element if and only if  $y'_n = y_n$ . We define the partitions  $\alpha_n^K, \eta_n^K, \tau_n^K$  of  $(K \setminus G)^{\mathbb{N}+1}$  similarly. We let  $\alpha_n^K(y)$  denote the partition element of  $\alpha_n^K$  that contains  $y$  (and similar notation holds for the other partitions).

Given partitions  $\alpha, \beta$  of a probability space  $(X, \kappa)$ , the entropy of  $\alpha$  relative to  $\beta$  is:

$$H(\alpha|\beta) := - \int \log(\kappa(\alpha(x)|\beta(x))) d\kappa(x)$$

where  $\alpha(x)$  denotes the partition element of  $\alpha$  containing  $x$  and  $\kappa(\alpha(x)|\beta(x)) = \frac{\kappa(\alpha(x) \cap \beta(x))}{\kappa(\beta(x))}$ .

**Lemma 3.2.** *For any  $K \in \text{Sub}_G$ ,*

$$\int H(\alpha_1^K|\eta_n^K) d\lambda(K) = \int H(\mu_K) - H(\mu_K^n) + H(\mu_K^{n-1}) d\lambda(K).$$

*Proof.* By definition, for any  $y \in (K \setminus G)^{\mathbb{N}+1}$ ,

$$\begin{aligned} \mathbb{P}_K(\alpha_1^K(y)|\eta_n^K(y)) &= \frac{\mathbb{P}_K(\{y' \in (K \setminus G)^{\mathbb{N}+1} : y'_1 = y_1, y'_i = y_i \forall i \geq n\})}{\mathbb{P}_K(\{y' \in (K \setminus G)^{\mathbb{N}+1} : y'_i = y_i \forall i \geq n\})} \\ &= \frac{\mathbb{P}_K(\{y' \in (K \setminus G)^{\mathbb{N}+1} : y'_1 = y_1, y'_n = y_n\})}{\mathbb{P}_K(\{y' \in (K \setminus G)^{\mathbb{N}+1} : y'_n = y_n\})} \\ &= \frac{\mu_K(y_1) \mathbb{P}_K(\tau_n^K(y)|\alpha_1^K(y))}{\mu_K(y_n)}. \end{aligned}$$



Note

$$\mathbb{P}_K(\tau_n^K(y)|\alpha_1^K(y)) = \mu_{y_1}^{n-1}(y_n).$$

We now have:

$$\mathbb{P}_K(\alpha_1^K(y)|\eta_n^K(y)) = \frac{\mu_K(y_1)\mu_{y_1}^{n-1}(y_n)}{\mu_K^n(y_n)}. \quad (1)$$

Therefore,

$$H(\alpha_1^K|\eta_n^K) = - \int \log(\mathbb{P}_K(\alpha_1^K(y)|\eta_n^K(y))) d\mathbb{P}_K(y) \quad (2)$$

$$= H(\mu_K) - H(\mu_K^n) + \sum_{g \in G} \mu(g) H(\mu_{Kg}^{n-1}). \quad (3)$$

Since  $\lambda$  is conjugation-invariant and  $H(\mu_{Kg}^{n-1}) = H(\mu_{g^{-1}Kg}^{n-1})$ ,

$$\int H(\mu_{Kg}^{n-1}) d\lambda(K) = \int H(\mu_{g^{-1}Kg}^{n-1}) d\lambda(K) = \int H(\mu_K^{n-1}) d\lambda(K).$$

So (3) implies

$$\int H(\alpha_1^K|\eta_n^K) d\lambda(K) = \int H(\mu_K) - H(\mu_K^n) + H(\mu_K^{n-1}) d\lambda(K).$$

□

**Lemma 3.3.** *The sequence  $\int H(\mu_K^n) - H(\mu_K^{n-1}) d\lambda(K)$  is monotone decreasing in  $n$ . Therefore,*

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \int H(\mu_K^n) d\lambda(K) &= \inf_{n \rightarrow \infty} \frac{1}{n} \int H(\mu_K^n) d\lambda(K) \\ &= \lim_{n \rightarrow \infty} \int H(\mu_K^n) - H(\mu_K^{n-1}) d\lambda(K) = \inf_{n \rightarrow \infty} \int H(\mu_K^n) - H(\mu_K^{n-1}) d\lambda(K) \\ &= \int H(\mu_K) - H(\alpha_1^K|\eta^K) d\lambda(K). \end{aligned}$$

*Proof.* Since  $\eta_{n-1}^K$  refines  $\eta_n^K$  we have  $H(\alpha_1^K|\eta_n^K) \geq H(\alpha_1^K|\eta_{n-1}^K)$ . So the previous lemma implies  $\int H(\mu_K^n) - H(\mu_K^{n-1}) d\lambda(K)$  is monotone decreasing in  $n$ . It is also bounded by  $H(\mu)$ . So,

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \int H(\mu_K^n) d\lambda(K) &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{m=1}^n \int H(\mu_K^m) - H(\mu_K^{m-1}) d\lambda(K) \\ &= \lim_{n \rightarrow \infty} \int H(\mu_K^n) - H(\mu_K^{n-1}) d\lambda(K) \\ &= \lim_{n \rightarrow \infty} \int H(\mu_K) - H(\alpha_1^K|\eta_n^K) d\lambda(K) \\ &= \int H(\mu_K) - H(\alpha_1^K|\eta^K) d\lambda(K). \end{aligned}$$

where  $\eta^K$  is the limit of  $\eta_n^K$  (so a set  $E$  is in the  $\sigma$ -algebra generated by  $\eta^K$  iff for every  $n$  it is in the  $\sigma$ -algebra generated by  $\eta_n^K$ ). The third line follows from the previous lemma.

Because  $\int H(\mu_K^n) - H(\mu_K^{n-1}) d\lambda(K)$  is monotone decreasing, it follows that

$$\begin{aligned} \frac{1}{n} \int H(\mu_K^n) d\lambda(K) &= \frac{1}{n} \sum_{m=1}^n \int H(\mu_K^m) - H(\mu_K^{m-1}) d\lambda(K) \\ &\geq \int H(\mu_K^n) - H(\mu_K^{n-1}) d\lambda(K). \end{aligned}$$

Therefore,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \int H(\mu_K^n) d\lambda(K) = \inf_n \int H(\mu_K^n) - H(\mu_K^{n-1}) d\lambda(K) = \inf_n \frac{1}{n} \int H(\mu_K^n) d\lambda(K).$$

□

**Lemma 3.4.** *For any  $K \in \text{Sub}_G$ ,*

$$H(\mu_K) - H(\alpha_1^K | \eta^K) = \sum_{g \in G} \mu_K(Kg) \int \log \left( \frac{d\nu_{Kg}}{d\nu_K}(b) \right) d\nu_{Kg}(b).$$

*Proof.* For any Borel  $E \subset B_K$  and any  $y \in (K \setminus G)^{\mathbb{N}+1}$ ,

$$\mathbb{P}_K(\{y' \in (K \setminus G)^{\mathbb{N}+1} : \mathbf{bnd}(y') \in E\} | \alpha_1^K(y)) = \nu_{y_1}(E) = \int_E \frac{d\nu_{y_1}}{d\nu_K}(b) d\nu_K(b).$$

Therefore,

$$\mathbb{P}_K(\alpha_1^K(y) | \eta^K(y)) = \mathbb{P}_K(\alpha_1^K(y)) \frac{d\nu_{y_1}}{d\nu_K}(\mathbf{bnd}(y))$$

for  $\mathbb{P}_K$  a.e.  $y$ . We now have:

$$\begin{aligned} H(\mu_K) - H(\alpha_1^K | \eta^K) &= H(\mu_K) + \int \log(\mathbb{P}_K(\alpha_1^K(y) | \eta^K(y))) d\mathbb{P}_K(y) \\ &= H(\mu_K) + \int \log \left( \mathbb{P}_K(\alpha_1^K(y)) \frac{d\nu_{y_1}}{d\nu_K}(\mathbf{bnd}(y)) \right) d\mathbb{P}_K(y) \\ &= \int \log \left( \frac{d\nu_{y_1}}{d\nu_K}(\mathbf{bnd}(y)) \right) d\mathbb{P}_K(y) \\ &= \sum_{g \in G} \int_{\{y: y_1 = Kg\}} \log \left( \frac{d\nu_{Kg}}{d\nu_K}(\mathbf{bnd}(y)) \right) d\mathbb{P}_K(y) \\ &= \sum_{g \in G} \mu_K(Kg) \int \log \left( \frac{d\nu_{Kg}}{d\nu_K}(b) \right) d\nu_{Kg}(b). \end{aligned}$$

□

**Lemma 3.5.** For any  $K \in \text{Sub}_G$ , Borel set  $E \subset B_K$  and  $\gamma \in G$ ,

$$\nu_K(E) = \nu_{\gamma K}(\gamma E).$$

*Proof.* The proof is immediate. □

**Lemma 3.6.** For  $\gamma \in G$  and  $(K, \xi) \in B(\text{Sub}_G)$ ,

$$\frac{d\nu_\lambda \circ \gamma^{-1}}{d\nu_\lambda}(K, \xi) = \frac{d\nu_{K\gamma}}{d\nu_K}(\xi).$$

*Proof.* By abuse of notation, we may regard  $B(\text{Sub}_G)$  as the disjoint union of sets  $B_K$  for  $K \in \text{Sub}_G$ . So for any Borel  $E \subset B(\text{Sub}_G)$ ,

$$\nu_\lambda \circ \gamma^{-1}(E) = \nu_\lambda(\gamma^{-1}E) = \int \nu_K(\gamma^{-1}E \cap B_K) d\lambda(K).$$

By Lemma 3.5,

$$\nu_K(\gamma^{-1}E \cap B_K) = \nu_{\gamma K}(E \cap B_{K\gamma}).$$

So

$$\nu_\lambda \circ \gamma^{-1}(E) = \int \nu_{\gamma K}(E \cap B_{K\gamma}) d\lambda(K).$$

Make the change of variable  $L = K\gamma$  and use the conjugation-invariance of  $\lambda$  to obtain

$$\nu_\lambda \circ \gamma^{-1}(E) = \int \nu_{L\gamma}(E \cap B_L) d\lambda(L).$$

In other words,

$$\nu_\lambda \circ \gamma^{-1}(E) = \iint \frac{d\nu_{K\gamma}}{d\nu_K}(\xi) 1_E(K, \xi) d\nu_K(\xi) d\lambda(K).$$

This implies the lemma. □

*Proof of Theorem 3.1.* By Lemmas 3.3, 3.4 and 3.6,

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \int H(\mu_K^n) d\lambda(K) &= \int H(\mu_K) - H(\alpha_1^K | \eta^K) d\lambda(K) \\ &= \int \sum_{g \in G} \mu_K(Kg) \int \log \left( \frac{d\nu_{Kg}}{d\nu_K}(b) \right) d\nu_{Kg}(b) d\lambda(K) \\ &= \sum_{g \in G} \mu(g) \iint \log \left( \frac{d\nu_{Kg}}{d\nu_K}(b) \right) \frac{d\nu_{Kg}}{d\nu_K}(b) d\nu_K(b) d\lambda(K) \\ &= \sum_{g \in G} \mu(g) \iint \log \left( \frac{d\nu_\lambda \circ g^{-1}}{d\nu_\lambda}(K, b) \right) \frac{d\nu_\lambda \circ g^{-1}}{d\nu_\lambda}(K, b) d\nu_\lambda(K, b). \end{aligned}$$

The cocycle identity for the Radon-Nikodym derivative implies

$$\frac{d\nu_\lambda \circ g^{-1}}{d\nu_\lambda}(K, b) = \frac{d\nu_\lambda}{d\nu_\lambda \circ g}(g^{-1}(K, b)).$$

By definition, we also have

$$\frac{d\nu_\lambda \circ g^{-1}}{d\nu_\lambda}(K, b) \, d\nu_\lambda(K, b) = d\nu_\lambda \circ g^{-1}(K, b).$$

Therefore,

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \int H(\mu_K^n) \, d\lambda(K) &= - \sum_{g \in G} \mu(g) \iint \log \left( \frac{d\nu_\lambda \circ g}{d\nu_\lambda}(g^{-1} \cdot (K, b)) \right) d\nu_\lambda(g^{-1} \cdot (K, b)) \\ &= - \sum_{g \in G} \mu(g) \iint \log \left( \frac{d\nu_\lambda \circ g}{d\nu_\lambda}(K, b) \right) d\nu_\lambda(K, b) \\ &= h_\mu(B(\text{Sub}_G), \nu_\lambda). \end{aligned}$$

The other equalities follow from Lemma 3.3. □

## 4 Results for the free group

For the sake of simplicity, we specialize to the case  $G = \langle a, b \rangle$ , the rank 2 free group although all the constructions easily generalize to any finitely generated free group.

Let  $\text{Schreier}(K \backslash G) = (V_K, E_K)$  be the Schreier coset graph of  $K \backslash G$ . The vertex set is  $V_K := K \backslash G$ . For each  $Kg \in K \backslash G$  there are two directed labeled edges in the edge set, denoted by  $E_K$ . These are  $(Kg, Kga)$  which is labeled  $a$ , and  $(Kg, Kgb)$  which is labeled  $b$ . It is possible that  $Kga = Kgb$  in which case there are two different edges from  $Kg$  to  $Kga = Kgb$ .

We say that  $K \backslash G$  is *tree-like* if for every  $Kg, Kg' \in K \backslash G$  there is a unique sequence of vertices  $Kg = Kg_1, Kg_2, \dots, Kg_n = Kg'$  such that  $Kg_i$  is adjacent to  $Kg_{i+1}$  for  $1 \leq i < n$  and  $Kg_{i-1} \neq Kg_{i+1}$  for any  $1 < i < n$ . This does not mean that there is a unique path in the Schreier coset graph of  $K \backslash G$  because it is possible, for example, that  $Kga = Kgb$  for some coset  $Kg$ . Equivalently,  $K \backslash G$  is treelike if it does not contain simple circuits of length greater than 2.

Let  $\text{Tree}_G \subset \text{Sub}_G$  be the set of all subgroups  $K \in \text{Sub}_G$  such that  $K \backslash G$  is tree-like. This is a closed  $G$ -invariant subspace. Let  $\mathcal{M}(\text{Tree}_G) \subset \mathcal{M}(\text{Sub}_G)$  denote those measures with support contained in  $\text{Tree}_G$ .

Let  $\{X_i\}_{i=1}^\infty$  be i.i.d. random variables in  $G$  with law  $\mu$  (where  $\mu$  is the uniform probability measure on  $\{a, a^{-1}, b, b^{-1}\}$ ). For  $K \in \text{Sub}_G$ , let  $R_n(\mu, K)$  be the probability that  $KX_1 \cdots X_n = K$  and let  $R_{\geq n}(\mu, K)$  be the probability that  $KX_1 \cdots X_m = K$  for some  $m \geq n$ . A subset  $\mathcal{N} \subset \mathcal{M}(\text{Sub}_G)$  has *controlled return-time probabilities* if

$$\lim_{\epsilon \rightarrow 0^+} \limsup_{n \rightarrow \infty} \eta \left( \{K \in \text{Sub}_G : R_{\geq n}(\mu, K) \geq \epsilon\} \right) = 0.$$

The next result plays a key role in the proof of Theorem 1.1. It is proven in the next subsection.

**Theorem 4.1.** *If  $\mathcal{N} \subset \mathcal{M}(\text{Tree}_G)$  is a set of measures with controlled return-time probabilities then the entropy function  $\lambda \in \mathcal{N} \mapsto h_\mu(B(\text{Sub}_G), \nu_\lambda)$  is continuous on  $\mathcal{N}$  with respect to the weak\* topology.*

## 4.1 A continuity criterion

If  $K \in \text{Tree}_G$  (so  $K \setminus G$  is tree-like), then for each  $g \in G$ , let the *shadow* of  $Kg$ , denoted  $\text{Shd}(Kg)$ , be the set of all cosets  $K\gamma \in K \setminus G$  so that every path in  $\text{Schreier}(K \setminus G)$  from  $K$  to  $K\gamma$  passes through  $Kg$ . Let  $\text{Shd}_{\mathbb{N}}(Kg)$  be the set of all sequences  $(Kg_1, Kg_2, \dots) \in (K \setminus G)^{\mathbb{N}}$  that are eventually in  $\text{Shd}(Kg)$  in the sense that there exists an  $N$  so that if  $n \geq N$  then  $Kg_n \in \text{Shd}(Kg)$ . Let  $B_{Kg} := \pi_K(\text{Shd}_{\mathbb{N}}(Kg))$  be the projection of  $\text{Shd}_{\mathbb{N}}(Kg)$  to the boundary  $B_K$ .

**Lemma 4.2.** *Let  $\lambda \in \mathcal{M}(\text{Tree}_G)$  and  $s, t \in \{a, b, a^{-1}, b^{-1}\}$ . Then for  $\nu_\lambda$ -a.e.  $(K, \xi)$ , if  $\xi \in B_{Ks}$  then*

$$\frac{d\nu_\lambda \circ t}{d\nu_\lambda}(K, \xi) = \frac{\nu_{Kt^{-1}}(B_{Ks})}{\nu_K(B_{Ks})}.$$

*Proof.* By Lemma 3.6,

$$\frac{d\nu_\lambda \circ t}{d\nu_\lambda}(K, \xi) = \frac{d\nu_{Kt^{-1}}}{d\nu_K}(\xi).$$

Let  $\{X_n(Kg) : Kg \in K \setminus G, n \geq 1\}$  be an i.i.d. family of random variables with law  $\mu$ . Let  $\{Z_n(Kg)\}_{n=1}^\infty$  be the random walk:

$$Z_n(Kg) := KgX_1(Kg)X_2(Kg) \cdots X_n(Kg)$$

and  $Z_0(Kg) := Kg$ .

Recall that  $\mathbf{bnd}_K$  denotes the projection from the space of sequences  $(K \setminus G)^{\mathbb{N}+1}$  to the boundary  $B_K$ . Let  $\zeta(Kg) = \mathbf{bnd}_K(\{Z_n(Kg)\}_{n=0}^\infty)$ .

Suppose that  $t^{-1} \neq s$ . Then any path in  $\text{Schreier}(K \setminus G)$  from  $Kt^{-1}$  whose projection lies in  $B_{Ks}$  necessarily passes through  $K$ . So for any Borel  $E \subset B_{Ks}$  the probability that  $\zeta(Kt^{-1}) \in E$  is

$$\begin{aligned} & \nu_{Kt^{-1}}(E) = \Pr(\zeta(Kt^{-1}) \in E) \\ &= \sum_{n=0}^{\infty} \Pr(Z_n(Kt^{-1}) = K, Z_m(Kt^{-1}) \neq K \ \forall m > n) \cdot \Pr(\zeta(K) \in E, Z_t(K) \neq K \ \forall t > 0) \\ &= \Pr(\exists n \text{ s.t. } Z_n(Kt^{-1}) = K) \cdot \Pr(\zeta(K) \in E \text{ and } Z_t(K) \neq K \ \forall t > 0) \\ &= \Pr(\exists n \text{ s.t. } Z_n(Kt^{-1}) = K) \cdot \frac{\Pr(\zeta(K) \in E)}{\sum_{t=0}^{\infty} \Pr(Z_t(K) = K)} \\ &= \nu_K(E) \frac{\Pr(\exists n \text{ s.t. } Z_n(Kt^{-1}) = K)}{\sum_{t=0}^{\infty} \Pr(Z_t(K) = K)}. \end{aligned}$$

Since this is true for every  $E \subset B_{Ks}$  it follows that

$$\frac{d\nu_{Kt^{-1}}}{d\nu_K}(\xi) = \frac{\Pr(\exists n \text{ s.t. } Z_n(Kt^{-1}) = K)}{\sum_{t=0}^{\infty} \Pr(Z_t(K) = K)}.$$

In particular,  $\frac{d\nu_{Kt^{-1}}}{d\nu_K}(\xi) = \frac{\nu_{Kt^{-1}}(E)}{\nu_K(E)}$  for every measurable  $E \subset B_{Ks}$  with positive measure. This proves the lemma in the case  $Kt^{-1} \neq Ks$ .

Suppose now that  $Kt^{-1} = Ks$ . Then any path in  $\text{Schreier}(K \setminus G)$  from  $K$  which projects into  $B_{Ks}$  must pass through  $Ks$ . So for any subset  $E \subset B_{Ks}$ ,

$$\begin{aligned} \nu_K(E) &= \Pr(\zeta(K) \in E) \\ &= \sum_{n=0}^{\infty} \Pr(Z_n(K) = Ks, Z_m(K) \neq Ks \ \forall m > n) \cdot \Pr(\zeta(Ks) \in E, Z_r(Ks) \neq Ks \ \forall r > 0) \\ &= \Pr(\exists n \text{ such that } Z_n(K) = Ks) \cdot \Pr(\zeta(Ks) \in E, Z_r(Ks) \neq Ks \ \forall r > 0) \\ &= \Pr(\exists n \text{ such that } Z_n(K) = Ks) \frac{\Pr(\zeta(Ks) \in E)}{\sum_{n=0}^{\infty} \Pr(Z_n(Ks) = Ks)} \\ &= \nu_{Ks}(E) \frac{\Pr(\exists n \text{ such that } Z_n = Ks)}{\sum_{n=0}^{\infty} \Pr(Z_n(Ks) = Ks)}. \end{aligned}$$

Since this is true for every  $E \subset B_{Ks}$  it follows that

$$\frac{d\nu_{Kt^{-1}}}{d\nu_K}(\xi) = \frac{d\nu_{Ks}}{d\nu_K}(\xi) = \frac{\sum_{n=0}^{\infty} \Pr(Z_n(Ks) = Ks)}{\Pr(\exists n \text{ such that } Z_n = Ks)}.$$

In particular,  $\frac{d\nu_{Kt^{-1}}}{d\nu_K}(\xi) = \frac{\nu_{Kt^{-1}}(E)}{\nu_K(E)}$  for every Borel  $E \subset B_{Ks}$  with positive measure. This proves the lemma in the case  $Kt^{-1} = Ks$ .  $\square$

**Lemma 4.3.** *For any  $t, s \in \{a, b, a^{-1}, b^{-1}\}$  and  $K \in \text{Tree}_G$ ,*

$$1/4 \leq \frac{\nu_{Kt^{-1}}(B_{Ks})}{\nu_K(B_{Ks})} \leq 4.$$

*Proof.* For  $n \geq 0$  and  $g \in G$ , define  $Z_n(Kg)$  and  $\zeta(Kg)$  as in the proof of the previous lemma. Let  $s \in \{a, b, a^{-1}, b^{-1}\}$  be such that  $\xi \in B_{Ks}$ . Then

$$\begin{aligned} \nu_{Kt^{-1}}(B_{Ks}) &= \Pr(\zeta(Kt^{-1}) \in B_{Ks}) \\ &\geq \Pr(Z_1(Kt^{-1}) = K) \cdot \Pr(\zeta(K) \in B_{Ks}) \\ &\geq \nu_K(B_{Ks})/4. \end{aligned}$$

The other inequality is similar.  $\square$

*Proof of Theorem 4.1.* Let  $S = \{a, b, a^{-1}, b^{-1}\}$ . By Lemma 4.2, for any  $\lambda \in \mathcal{M}(\text{Tree}_G)$ ,

$$\begin{aligned} h_\mu(B(\text{Sub}_G), \nu_\lambda) &= - \iint \log \frac{d\nu_\lambda \circ g}{d\nu_\lambda}(K, \xi) d\nu_\lambda(K, \xi) d\mu(g) \\ &= - \iint \sum_{s \in S} \nu_K(B_{Ks}) \log \frac{\nu_{Kg^{-1}}(B_{Ks})}{\nu_K(B_{Ks})} d\lambda(K) d\mu(g) \\ &= - \int \sum_{t \in S} \sum_{s \in S} \mu(t) \nu_K(B_{Ks}) \log \frac{\nu_{Kt^{-1}}(B_{Ks})}{\nu_K(B_{Ks})} d\lambda(K). \end{aligned}$$

By the previous lemma there is a constant  $C > 0$  so that

$$\left| \nu_K(B_{Ks}) \log \frac{\nu_{Kt^{-1}}(B_{Ks})}{\nu_K(B_{Ks})} \right| \leq C$$

for all  $K \in \text{Tree}_G$  and  $s, t \in \{e, a, b, a^{-1}, b^{-1}\}$ . For  $x, y \in [0, 1]$  let

$$F(x, y) := \begin{cases} -C & -x \log \frac{y}{x} \leq -C \\ -x \log \frac{y}{x} & -C < -x \log \frac{y}{x} < C \\ C & -x \log \frac{y}{x} \geq C \end{cases}$$

Also for  $s, t \in \{a, b, a^{-1}, b^{-1}, e\}$  and  $K \in \text{Sub}_G$ , let

$$\rho(K, s, t) := \mu(t) F(\nu_K(B_{Ks}), \nu_{Kt^{-1}}(B_{Ks})).$$

So the previous equation implies

$$h_\mu(B(\text{Sub}_G), \nu_\lambda) = \sum_{t \in S} \sum_{s \in S} \int \rho(K, s, t) d\lambda(K). \quad (4)$$

Define  $Z_n(Kg)$  and  $\zeta(Kg)$  as in Lemma 4.2. For  $n, \epsilon \geq 0$ ,  $t, s \in \{a, b, a^{-1}, b^{-1}, e\}$  and  $K \in \text{Sub}_G$  let

$$\rho_n(K, s, t) := \mu(t) F(\Pr(Z_n(K) \in \text{Shd}(Ks)), \Pr(Z_n(Kt^{-1}) \in \text{Shd}(Ks))).$$

Note that  $\rho_n(K, s, t)$  varies continuously with  $K \in \text{Tree}_G$  (for fixed  $s, t, n$ ). Since we are using the weak\* topology on  $\mathcal{N} \subset \mathcal{M}(\text{Sub}_G)$ , it suffices to show that

$$\lim_{n \rightarrow \infty} \sup_{\lambda \in \mathcal{N}} \left| h_\mu(B(\text{Sub}_G), \nu_\lambda) - \sum_{t \in S} \sum_{s \in S} \int \rho_n(K, s, t) d\lambda(K) \right| = 0.$$

By (4) it suffices to prove

$$\lim_{n \rightarrow \infty} \sup_{\lambda \in \mathcal{N}} \sum_{t \in S} \sum_{s \in S} \int |\rho(K, s, t) - \rho_n(K, s, t)| d\lambda(K) = 0.$$

Let

$$X_{n,s,t}(\epsilon) := \{K \in \text{Sub}_G : |\rho(K, s, t) - \rho_n(K, s, t)| < \epsilon\}.$$

Because  $|\rho(K, s, t) - \rho_n(K, s, t)|$  is bounded by  $2C$ , it suffices to show that

$$\lim_{\epsilon \searrow 0} \lim_{n \rightarrow \infty} \sup_{\lambda \in \mathcal{N}} \lambda(X_{n,s,t}(\epsilon)) = 1.$$

For  $\delta > 0$ , let

$$Y_{n,s,t}(\delta) := \{K \in \text{Sub}_G : |\nu_{Kt^{-1}}(B_{Ks}) - \Pr(Z_n(Kt^{-1}) \in \text{Shd}(Ks))| < \delta\}.$$

Because  $F$  is uniformly continuous on  $[0, 1] \times [0, 1]$ , it suffices to prove that

$$\lim_{\delta \searrow 0} \lim_{n \rightarrow \infty} \sup_{\lambda \in \mathcal{N}} \lambda(Y_{n,s,t}(\delta)) = 1 \tag{5}$$

for every  $s, t$ .

Because  $\mathcal{N}$  has controlled return-time probabilities, for any  $\delta > 0$  there exists an  $N = N(\delta)$  such that  $n \geq N$  implies

$$\lambda(\{K \in \text{Sub}_G : R_{\geq n}(\mu, K) \geq \delta\}) < \delta \quad \forall \lambda \in \mathcal{N}.$$

Equivalently,

$$\lambda(\{K \in \text{Sub}_G : \Pr(Z_n(K) \neq K, \forall n \geq N) \geq 1 - \delta\}) \geq 1 - \delta \quad \forall \lambda \in \mathcal{N}.$$

Because the support  $\mu$  is contained in  $\{a, b, a^{-1}, b^{-1}\}$  if  $Z_n(K) \neq K$  for any  $n \geq N$  and  $Z_n(K) \in \text{Shd}(Ks)$  (for some  $s \in \{a, b, a^{-1}, b^{-1}\}$ ) then  $\zeta(K) \in B_{Ks}$ . Therefore, the equation above implies that for any  $s \in \{a, b, a^{-1}, b^{-1}\}$  and any  $n \geq N$ ,

$$\lambda(\{K \in \text{Sub}_G : |\Pr(Z_n(K) \in \text{Shd}(Ks)) - \Pr(\zeta(K) \in B_{Ks})| \leq \delta\}) \geq 1 - \delta, \quad \forall \lambda \in \mathcal{N}.$$

Since  $\Pr(\zeta(K) \in B_{Ks}) = \nu_K(B_{Ks})$ , this equation is equivalent to

$$\lambda(\{K \in \text{Sub}_G : |\Pr(Z_n(K) \in \text{Shd}(Ks)) - \nu_K(B_{Ks})| \leq \delta\}) \geq 1 - \delta, \quad \forall \lambda \in \mathcal{N}.$$

This implies equation (5) for  $t = e$ . The other cases are similar. □

## 4.2 A covering space construction

For  $K \in \text{Tree}_G$ , let  $X_K$  be the 2-complex whose 1-skeleton is the right-Schreier coset graph of  $K$  (with respect to  $G$ ) and whose 2-cells are all possible nonagons and bigons. More precisely, for every loop in the Schreier coset graph, there is a 2-cell whose boundary is that



loop and if  $e_1, e_2$  are two edges with the same endpoints, then there is a 2-cell with boundary  $e_1 \cup e_2$ . Because  $K \setminus G$  is tree-like,  $X_K$  is simple-connected.

If  $c$  is a 2-cell of  $X_K$  and  $g \in G$ , then we let  $gc$  be the corresponding 2-cell of  $X_{gKg^{-1}}$ . For example, if  $c$  bounds a loop based at the vertex  $Kh \in X_K^{(0)}$  then  $gc$  bounds a loop based at the vertex  $gKh \in X_{gKg^{-1}}^{(0)}$ . If  $c$  bounds a pair of edges  $(Kh, Khs), (Kh, Kht)$  (for some  $t, s \in S := \{a, b, a^{-1}, b^{-1}\}$ ) with the same endpoints, then  $gc$  bounds the pair of edges  $(gKh, gKhs), (gKh, gKht)$ .

Let  $\widetilde{\text{Tree}}_G$  be the set of all pairs  $(K, \omega)$  where  $K \in \text{Tree}_G$  and  $\omega \subset X_K^{(2)}$  is a collection of 2-cells of  $X_K$ .  $G$  acts on this space by  $g(K, \omega) = (gKg^{-1}, g\omega)$  where  $g\omega = \{gc : c \in \omega\}$ .

There is a natural topology on  $\widetilde{\text{Tree}}_G$ . To explain, let  $B_n(X_K)$  denote the ball of radius  $n$  centered at  $K$  in  $X_K$ , considered as a 2-complex with oriented edges labeled by the generating set  $S$ . For each integer  $n \geq 1$  and  $(K, \omega) \in \widetilde{\text{Tree}}_G$ , let  $Nbhd_n(K, \omega)$  be the set of all  $(K', \omega') \in \widetilde{\text{Tree}}_G$  such that there is an isomorphism  $\phi : B_n(X_K) \rightarrow B_n(X_{K'})$  which preserves labels and directions on the edges and also maps  $B_n(X_K) \cap \omega$  bijectively onto  $B_n(X_{K'}) \cap \omega'$ . We obtain a topology on  $\widetilde{\text{Tree}}_G$  by declaring that each  $Nbhd_n(K, \omega)$  is clopen.

Let  $(K, \omega) \in \widetilde{\text{Tree}}_G$ . If  $\omega$  is nonempty then  $X_K \setminus \omega$  is not simply-connected. So let  $\pi_{K, \omega} : U_{K, \omega} \rightarrow X_K \setminus \omega$  be the universal cover. Observe that there is a natural right-action of  $G$  on  $U_{K, \omega}$  obtained by path continuation. Choose a vertex  $u_{K, \omega} \in U_{K, \omega}^{(0)}$  which projects to  $K \setminus G$  and let  $S_{K, \omega}$  be the stabilizer  $S_{K, \omega} := \{g \in G : u_{K, \omega}g = u_{K, \omega}\}$ . Because  $S_{K, \omega} \setminus G$  is naturally identified with the 1-skeleton of  $U_{K, \omega}$ , it follows that  $S_{K, \omega} \setminus G$  is tree-like. Also, observe that  $S_{K, \omega}$  does not depend on the choice of  $u_{K, \omega}$ . Indeed, it is the subgroup of  $K$  generated by all elements of the form

1.  $gsg^{-1}$  for every 2-cell **not** in  $\omega$  which bounds a loop based at  $Kg$  labeled  $s \in \{a, b, a^{-1}, b^{-1}\}$ ;
2.  $gs_1s_2^{-1}g^{-1}$  for every 2-cell **not** in  $\omega$  which bounds a bigon whose edges are labeled  $s_1, s_2 \in \{a, b, a^{-1}, b^{-1}\}$  and are directed from  $Kg$  to  $Kgs_1 = Kgs_2$ .

The fundamental group of  $X_K \setminus \omega$  is  $K/S_{K, \omega}$ . In particular,  $S_{K, \omega}$  is normal in  $K$ .

The map from  $\Psi : \widetilde{\text{Tree}}_G \rightarrow \text{Tree}_G$  defined by  $\Psi(K, \omega) = S_{K, \omega}$  is  $G$ -equivariant. Therefore, if  $\tilde{\eta}$  is a  $G$ -invariant ergodic probability measure on  $\widetilde{\text{Tree}}_G$ , then,  $\Psi_*\tilde{\eta}$  is conjugation-invariant and ergodic. Also,  $\Psi$  is continuous so  $\Psi_* : \mathcal{M}(\widetilde{\text{Tree}}_G) \rightarrow \mathcal{M}(\text{Tree}_G)$  is continuous in the weak\* topology where  $\mathcal{M}(\widetilde{\text{Tree}}_G)$  denotes the space of  $G$ -invariant Borel probability measures on  $\widetilde{\text{Tree}}_G$ .

**Lemma 4.4.** *For any subgroup  $K < G$ , let  $\overline{R}(K)$  be the expected number of returns of the random walk on  $K \setminus G$  to  $K$ . That is,*

$$\overline{R}(K) = \int \#\{n : Kg_n = K\} d\mathbb{P}(g_0, g_1, \dots).$$

*For  $\lambda \in \mathcal{M}(\text{Sub}_G)$ , let  $\overline{R}(\lambda) = \int \overline{R}(K) d\lambda(K)$ . If  $\tilde{\lambda}$  is a  $G$ -equivariant measure on  $\widetilde{\text{Tree}}_G$  which projects to  $\lambda \in \mathcal{M}(G)$ , then  $\overline{R}(\lambda) \geq \overline{R}(\Psi_*\tilde{\lambda})$ .*

*Proof.* For each  $g_n \in G$ ,  $(K, \omega) \in \widetilde{\text{Tree}_G}$ , if  $S_{K, \omega} g_n = S_{K, \omega}$  then  $K g_n = K$  (because  $S_{K, \omega} < K$ ). So  $\overline{R}(S_{K, \omega}) \leq \overline{R}(K)$  and

$$\overline{R}(\Psi_* \tilde{\lambda}) = \int \overline{R}(S_{K, \omega}) d\tilde{\lambda}(K, \omega) \leq \int \overline{R}(K) d\lambda = \overline{R}(\lambda).$$

□

**Notation 1.** To simplify notation, for any  $\lambda \in \mathcal{M}(\text{Sub}_G)$ , let  $h_\mu(\lambda) := h_\mu(B(\text{Sub}_G), \nu_\lambda)$ .

**Corollary 4.5.** For  $\eta \in \mathcal{M}(\text{Tree}_G)$ , let  $\mathcal{M}_\eta(\widetilde{\text{Tree}_G})$  be the space of all  $G$ -invariant Borel probability measures on  $\widetilde{\text{Tree}_G}$  which project to  $\eta$ . This is a compact convex space under the weak\* topology. Moreover, if  $\overline{R}(\eta) < \infty$  then the map which sends  $\lambda \in \widetilde{\text{Tree}_G}$  to  $h_\mu(\Psi_* \lambda)$  is continuous on  $\mathcal{M}_\eta(\widetilde{\text{Tree}_G})$ .

*Proof.* By the previous lemma,  $\infty > \overline{R}(\eta) \geq \overline{R}(\Psi_* \lambda)$  for all  $\lambda \in \mathcal{M}_\eta(\widetilde{\text{Tree}_G})$ . Therefore,  $\mathcal{M}_\eta(\widetilde{\text{Tree}_G})$  has controlled return-time probabilities. So the corollary follows from Theorem 4.1. □

In order to prove Theorem 1.1, it now suffices to show there exists a sequence  $\{\eta_n\}_{n=1}^\infty \subset \mathcal{M}(\text{Tree}_G)$  and for every  $n$ , a continuous 1-parameter family  $\{\eta_{n,p} : 0 \leq p \leq 1\} \subset \mathcal{M}_{\eta_n}(\widetilde{\text{Tree}_G})$  of ergodic measures such that  $\overline{R}(\eta_n) < \infty$  for all  $n$ ,  $\lim_{n \rightarrow \infty} h_\mu(\eta_n) = 0$ ,  $\Psi_*(\eta_{n,0}) = \eta_n$  and  $\Psi_*(\eta_{n,1})$  is the Dirac measure on the trivial subgroup. This is accomplished in the next section.

### 4.3 Paths of IRS's

For each integer  $n \geq 1$ , we define a subgroup  $K_n < G$  as follows (see figure 1 for an example). Let  $\phi : G \rightarrow \mathbb{Z}^2$  be the abelianization homomorphism.  $K_n$  is generated by all elements of the form  $ghg^{-1}$  where  $g \in \langle a^n, b^n \rangle$  and either  $h = a^k b^r a^{-k}$  for some  $1 \leq |k| \leq n-1$  and  $r \in \mathbb{Z}$  or  $h = b^k a^r b^{-k}$  for some  $1 \leq |k| \leq n-1$  and  $r \in \mathbb{Z}$ .

Note that there are only a finite number of  $G$ -conjugates of  $K_n$ . Indeed,

$$\mathcal{C}_n := \{a^{-i} K_n a^i, b^{-i} K_n b^i : 0 \leq i \leq n-1\}$$

is a complete set of conjugates. To see this, it suffices to show that for every group  $J \in \mathcal{C}_n$  and  $s \in \{a, b, a^{-1}, b^{-1}\}$ ,  $sJs^{-1} \in \mathcal{C}_n$ . For example, note that for any  $1 \leq i \leq n-1$ ,  $a^i b^{-1} a^{-i+1}, a^{i-1} b a^{-i} \in K_n$ . Therefore,

$$ba^{-i} K_n a^i b^{-1} = ba^{-i} (a^i b^{-1} a^{-i+1}) K_n (a^{i-1} b a^{-i}) a^i b^{-1} = a^{-i+1} K_n a^{i-1} \in \mathcal{C}_n.$$

The other cases are similar. Let  $\eta_n \in \mathcal{M}(G)$  be the measure uniformly distributed on  $\mathcal{C}_n$ , the set of  $K_n$  conjugates. From figure 1, it is apparent that  $K_n \setminus G$  is tree-like. So  $\eta_n \in \mathcal{M}(\text{Tree}_G)$ .

**Lemma 4.6.**  $\overline{R}(\eta_n)$  is finite.

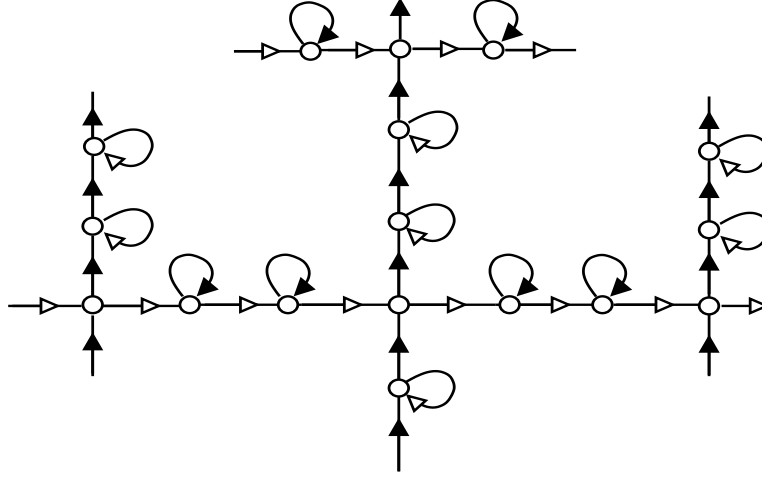


Figure 1: Part of the Schreier coset graph of  $K_3 \setminus G$ . The black arrows represent  $a$  and the white arrows represent  $b$ .

*Proof.* Let  $L_n = \cap_{g \in G} gK_n g^{-1}$  be the normal core of  $K_n$ . The discussion above implies that  $L_n$  has finite index in  $K_n$ . Also  $L_n$  is normal in  $G$ , so  $\delta_{L_n} \in \mathcal{M}(\text{Sub}_G)$  where  $\delta_{L_n}$  is the Dirac measure concentrated on  $\{L_n\}$ . We claim that  $G/L_n$  is non-amenable. Indeed,  $\langle a^n, b^n \rangle \cap K_n = \{1\}$  implies that  $G/L_n$  contains a free subgroup of rank 2 (the image of  $\langle a^n, b^n \rangle$ ). By Kesten's theorem, it follows that the spectral radius of the random walk on  $G/L_n$  is strictly less than 1 which implies  $\overline{R}(L_n)$  is finite.

Let  $K' < G$  be conjugate to  $K_n$ . Then the Schreier coset graph of  $K' \setminus G$  is finitely covered by the Schreier coset graph of  $L_n \setminus G$  with  $[K' : L_n]$  sheets. In particular, there are at most  $[K' : L_n]$  vertices of  $L_n \setminus G$  which map to  $K'$ . If  $(g_0, g_1, \dots)$  is a path in the Cayley graph of  $G$  and  $m$  is such that  $K'g_m = K'$  then, by the path lifting property of covering maps,  $L_n g_m$  is one of these  $[K' : L_n]$  vertices. Therefore,  $\overline{R}(K') \leq [K' : L_n] \overline{R}(L_n) = [K_n : L_n] \overline{R}(L_n)$ . Thus

$$\overline{R}(\eta_n) = |\mathcal{C}_n|^{-1} \sum_{K' \in \mathcal{C}_n} \overline{R}(K') \leq [K_n : L_n] \overline{R}(L_n) < \infty.$$

□

**Corollary 4.7.** *The map  $\lambda \in \mathcal{M}_{\eta_n}(\widetilde{\text{Tree}_G}) \mapsto h_\mu(\Psi_* \lambda)$  is continuous on  $\mathcal{M}_\eta(\widetilde{\text{Tree}_G})$ .*

*Proof.* This is implied by Corollary 4.5 and Lemma 4.6. □

**Lemma 4.8.** *Let  $\eta_{n,1} \in \mathcal{M}_\eta(\widetilde{\text{Tree}_G})$  be uniformly distributed on the  $G$ -orbit of  $(K_n, X_{K_n}^{(2)})$ . Then  $\Psi_* \eta_{n,1}$  is the trivial subgroup, so  $h_\mu(\Psi_* \eta_{n,1}) = h_{\max}(\mu)$ . Also  $\lim_{n \rightarrow \infty} h_\mu(\Psi_* \eta_n) = 0$ .*

*Proof.* The first claim is obvious. Note  $\eta_n$  converges in the weak\* topology to  $\delta_N$  (as  $n \rightarrow \infty$ ) where  $N$  is the kernel of the homomorphism  $G \rightarrow \mathbb{Z}$  given by  $a \rightarrow 1, b \rightarrow 1$  and  $\delta_N \in \mathcal{M}(\text{Sub}_G)$  is the Dirac probability measure concentrated on  $\{N\}$ . Because the random

walk on  $\mathbb{Z}$  has zero entropy,  $h_\mu(B(\text{Sub}_G), \nu_{\delta_N}) = 0$ . By Theorem 3.1,  $\lambda \mapsto h_\mu(B(\text{Sub}_G), \nu_\lambda)$  is an infimum of continuous functions and is therefore, upper semi-continuous. Therefore,

$$\limsup_{n \rightarrow \infty} h_\mu(\Psi_* \eta_{n,0}) \leq h_\mu(B(\text{Sub}_G), \nu_{\delta_N}) = 0.$$

□

Let  $\eta_{n,0}$  be the measure uniformly distributed on the  $G$ -orbit of  $(K_n, \emptyset)$ . Trivially,  $\Psi_*(\eta_{n,0}) = \eta_n$ . Because of the Lemma above and Corollary 4.7 to prove Theorem 1.1 it now suffices to show that for every  $n \geq 1$  there exists a continuous path of ergodic measures in  $\mathcal{M}_\eta(\widetilde{\text{Tree}_G})$  from  $\eta_{n,1}$  to  $\eta_{n,0}$ . We give two different proofs of this fact. The first is constructive. The second proof (in the next section) shows that in fact the entire space  $\mathcal{M}_{\eta_n}^e(\widetilde{\text{Tree}_G})$  of ergodic measures in  $\mathcal{M}_{\eta_n}(\widetilde{\text{Tree}_G})$  is pathwise connected.

*Proof of Theorem 1.1.* Let  $0 \leq p \leq 1$  and  $n \geq 1$  be an integer. Let  $K'$  be a uniformly random conjugate of  $K_n$ . Let  $\omega$  be the random element of  $X_K^{(2)}$  satisfying

- for every disjoint pair of finite sets  $Y, Z \subset X_K^{(2)}$ , the probability that  $Y \subset \omega$  and  $Z \cap \omega = \emptyset$  is  $p^{|Y|}(1-p)^{|Z|}$ .

Let  $\eta_{n,p}$  be the law of  $(K', \omega)$ . It is a  $G$ -invariant ergodic probability measure on  $\widetilde{\text{Tree}_G}$ . Also,  $p \mapsto \eta_{n,p}$  is continuous. So Corollary 4.7 implies  $p \mapsto h_\mu(\Psi_* \eta_{n,p})$  is continuous. By Lemma 4.8, for every  $t$  with  $h_\mu(\eta_n) \leq t \leq h_{\max}(\mu)$ , there is a  $p \in [0, 1]$  such that  $h_\mu(\Psi_* \eta_{n,p}) = t$ . Because  $\lim_{n \rightarrow \infty} h_\mu(\eta_n) = 0$ , this implies the theorem. □

It may interest the reader to know that the paths  $p \mapsto h_\mu(\Psi_* \eta_{n,p})$  are monotone increasing. This follows from the next lemma and corollary.

**Lemma 4.9.** *Let  $\rho$  be a Borel probability measure on  $\{(K_1, K_2) \in \text{Sub}_G \times \text{Sub}_G : K_1 < K_2\}$ . Suppose  $\rho$  is invariant under the diagonal action of  $G$  by conjugation. For  $i = 1, 2$ , let  $\rho_i$  be the projection of  $\rho$  onto the  $i$ -th coordinate. Then  $h_\mu(B(\text{Sub}_G), \nu_{\rho_1}) \geq h_\mu(B(\text{Sub}_G), \nu_{\rho_2})$ .*

*Proof.* Observe that if  $K_1 < K_2$  then  $H(\mu_{K_1}^n) \geq H(\mu_{K_2}^n)$  since the projection map  $K_1 \setminus G \rightarrow K_2 \setminus G_2$  maps  $\mu_{K_1}^n$  onto  $\mu_{K_2}^n$ . By Theorem 3.1,

$$\begin{aligned} h_\mu(B(\text{Sub}_G), \nu_{\rho_1}) &= \lim_{n \rightarrow \infty} \frac{1}{n} \int H(\mu_K^n) d\rho_1(K) = \lim_{n \rightarrow \infty} \frac{1}{n} \int H(\mu_{K_1}^n) d\rho(K_1, K_2) \\ &\geq \lim_{n \rightarrow \infty} \frac{1}{n} \int H(\mu_{K_2}^n) d\rho(K_1, K_2) = \lim_{n \rightarrow \infty} \frac{1}{n} \int H(\mu_K^n) d\rho_2(K) \\ &= h_\mu(B(\text{Sub}_G), \nu_{\rho_2}). \end{aligned}$$

□

**Corollary 4.10.** *The paths  $p \mapsto h_\mu(\Psi_* \eta_{n,p})$  are monotone increasing.*

*Proof.* Let  $K \in \text{Tree}_G$  be random with law  $\eta_n$ . Let  $\text{Leb}$  denote Lebesgue measure on  $[0, 1]$  and let  $x : X_K^{(2)} \rightarrow [0, 1]$  be random with law  $\text{Leb}^{X_K^{(2)}}$ . In other words, for each cell  $c \in X_K^{(2)}$ ,  $x(c)$  has law  $\text{Leb}$  and the variables  $\{x(c) : c \in X_K^{(2)}\}$  are independent.

Fix  $p, q$  with  $0 \leq p \leq q \leq 1$ . Let  $\omega_p = x^{-1}([0, p])$  and  $\omega_q = x^{-1}([0, q])$ . Let  $\rho$  be the law of the pair  $(S_{K, \omega_q}, S_{K, \omega_p})$  (where  $S_{K, \omega}$  is defined in §4.2). Clearly the projection of  $\rho$  onto its first factor is  $\Psi_* \eta_{n, q}$  and the projection onto its second factor is  $\Psi_* \eta_{n, p}$ . Because  $\omega_p \subset \omega_q$ , it follows that  $S_{K, \omega_q} < S_{K, \omega_p}$ . So the previous lemma implies  $h_\mu(\Psi_* \eta_{n, q}) \geq h_\mu(\Psi_* \eta_{n, p})$  as required.  $\square$

## 4.4 Entropies of boundaries of quotient groups

In this subsection, we prove Theorem 1.2.

**Lemma 4.11.** *If  $K < G$  has only finitely many conjugates,  $\eta \in \mathcal{M}(\text{Sub}_G)$  is the uniform measure on the set of conjugates of  $K$  and  $N = \cap_{g \in G} gKg^{-1}$  is the normal core of  $K$ , then  $h_\mu(B(\text{Sub}_G), \nu_\eta) = h_\mu(B_N, \nu_N)$ .*

*Proof.* By hypothesis,  $N$  has finite index in  $K$ . The projection map  $N \backslash G \rightarrow K \backslash G$  is  $[K : N]$ -to-1. Therefore,

$$H(\mu_K^n) \leq H(\mu_N^n) \leq H(\mu_K^n) + \log([K : N]).$$

A similar statement holds with  $K$  replaced by any of its conjugates  $gKg^{-1}$ . By Theorem 3.1,

$$\begin{aligned} h_\mu(B(\text{Sub}_G), \nu_\eta) &= \lim_{n \rightarrow \infty} \frac{1}{n} \int H(\mu_L^n) d\eta(L) \\ &\leq \lim_{n \rightarrow \infty} \frac{1}{n} H(\mu_N^n) = h_\mu(B_N, \nu_N) \\ &\leq \lim_{n \rightarrow \infty} \frac{1}{n} \int H(\mu_L^n) d\eta(L) + \frac{\log([K : N])}{n} \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \int H(\mu_L^n) d\eta(L) = h_\mu(B(\text{Sub}_G), \nu_\eta). \end{aligned}$$

$\square$

A  $G$ -invariant Borel probability measure  $\tilde{\eta}$  on  $\widetilde{\text{Tree}_G}$  is *periodic* if it has finite support.

**Lemma 4.12.** *Let  $\tilde{\eta}$  be an ergodic periodic measure on  $\widetilde{\text{Tree}_G}$ . Then there exists a normal subgroup  $N \triangleleft G$  such that  $h_\mu(\Psi_* \tilde{\eta}) = h_\mu(B_N, \nu_N)$ .*

*Proof.* Because  $\tilde{\eta}$  is ergodic and periodic, there exists an element  $(K, \omega) \in \widetilde{\text{Tree}_G}$  such that  $\tilde{\eta}$  equal the uniform probability measure on the  $G$ -orbit of  $(K, \omega)$ . Therefore, there is an element  $K' \in \text{Sub}_G$  such that  $\Psi_* \tilde{\eta}$  is the uniform probability measure on the (finite) set of conjugates of  $K'$ . Let  $N = \cap_{g \in G} gK'g^{-1}$ . By the previous lemma,  $h_\mu(\Psi_* \tilde{\eta}) = h_\mu(B_N, \nu_N)$ .  $\square$

**Lemma 4.13.** *Let  $\eta_n$  be the uniform probability measure on the set of conjugates of  $K_n$  as above. Let  $\mathcal{M}_{\eta_n}(\widetilde{\text{Tree}_G})$  be the space of all  $G$ -invariant Borel probability measures on  $\widetilde{\text{Tree}_G}$  that project to  $\eta_n$ . Let  $\mathcal{M}_{\eta_n}^p(\widetilde{\text{Tree}_G}) \subset \mathcal{M}_{\eta_n}(\widetilde{\text{Tree}_G})$  be the subspace of ergodic periodic measures. Then  $\mathcal{M}_{\eta_n}^p(\widetilde{\text{Tree}_G})$  is dense in  $\mathcal{M}_{\eta_n}(\widetilde{\text{Tree}_G})$ .*

*Proof.* Let  $\text{Loop}_n$  be the set of all elements of  $G$  of the form  $ga^i$  or  $gb^i$  for  $g \in \langle a^n, b^n \rangle$  and  $1 \leq i \leq n-1$ . Then  $\text{Loop}_n$  is naturally in bijection with the set of all loops in  $X_{K_n}$  which is naturally in bijection with the collection  $X_{K_n}^{(2)}$  of its 2-cells. Indeed, for every  $ga^i \in \text{Loop}_n$  there is a loop based at  $K_n ga^i$  with label  $b$  and for every  $gb^i \in \text{Loop}_n$  there is a loop based at  $K_n gb^i$  with label  $a$ . This is the bijection.

Fix  $n \geq 1$  and let  $\Omega$  be the set of all subsets of  $\{a^1, \dots, a^{n-1}, b^1, \dots, b^{n-1}\}$ . Let  $G$  act on  $\Omega^G$  by  $(g \cdot x)(f) = x(g^{-1}f)$ .

Let  $\Phi : \Omega^G \rightarrow \widetilde{\text{Tree}_G}$  be the map  $\Phi(x) = (K_n, \omega_x)$  where the 2-cell based at  $K_n ga^i$  (for  $g \in \langle a^n, b^n \rangle$  and  $1 \leq i \leq n-1$ ) is in  $\omega_x$  if and only if  $x(g)$  contains  $a^i$ . Similarly, the 2-cell based at  $K_n gb^i$  in  $\omega_x$  if and only if  $x(g)$  contains  $b^i$ . Note that  $\Phi$  is a homeomorphism onto the subspace  $\{(K_n, \omega) : \omega \subset X_{K_n}^{(2)}\}$ .

Let  $\phi : G \rightarrow \langle a^n, b^n \rangle$  be the isomorphism defined by  $\phi(a) = a^n, \phi(b) = b^n$ . The map  $\Phi$  is equivariant in the sense that

$$\Phi(gx) = \phi(g)\Phi(x).$$

Let  $\mathcal{M}_G(\Omega^G)$  be the space of all  $G$ -invariant Borel probability measures on  $\Omega^G$  and let  $\mathcal{M}_{K_n}(\widetilde{\text{Tree}_G})$  be the space of all  $\langle a^n, b^n \rangle$ -invariant Borel probability measures on the set  $\{(K_n, \omega) : \omega \subset X_{K_n}^{(2)}\}$ . Because  $\Phi$  is equivariant and a homeomorphism, it follows that  $\Phi_*$  is an affine isomorphism from  $\mathcal{M}_G(\Omega^G)$  to  $\mathcal{M}_{K_n}(\widetilde{\text{Tree}_G})$ .

Let  $\Lambda : \mathcal{M}_{\eta_n}(\widetilde{\text{Tree}_G}) \rightarrow \mathcal{M}_{K_n}(\widetilde{\text{Tree}_G})$  be the normalized restriction map. In other words,  $\Lambda(\tilde{\eta})$  is  $|\mathcal{C}_n|^{-1}$  times the restrictions of  $\tilde{\eta}$  to  $\{(K_n, \omega) : \omega \subset X_{K_n}^{(2)}\}$  where  $|\mathcal{C}_n|$  is the number of conjugates of  $K_n$ . This map is an affine isomorphism. Indeed, its inverse is given by

$$\Lambda^{-1}(\tilde{\eta}) = \frac{\tilde{\eta} + \sum_{i=1}^{n-1} a^i \tilde{\eta} + b^i \tilde{\eta}}{2n-1}$$

because  $\{K_n, a^i K_n a^{-i}, b^i K_n b^{-i} : 1 \leq i \leq n-1\}$  is a complete set of conjugates for  $K_n$ .

It follows that  $\Lambda^{-1}\Phi$  is an affine isomorphism from  $\mathcal{M}_G(\Omega^G)$  to  $\mathcal{M}_{\eta_n}(\widetilde{\text{Tree}_G})$ . In particular, it maps ergodic measures to ergodic measures. Because  $\Phi$  is equivariant and  $\Lambda$  is a restriction map,  $\Lambda^{-1}\Phi$  takes periodic measures to periodic measures (where a measure  $\lambda \in \mathcal{M}_G(\Omega^G)$  is periodic if it has finite support). Therefore, it suffices to prove that the subspace of ergodic periodic measures in  $\mathcal{M}_G(\Omega^G)$  is dense in  $\mathcal{M}_G(\Omega^G)$ . The next lemma shows how this result follows from [Ke12, Theorem 3.1] (it was proven earlier that the subspace of periodic measures is dense [Bo03]).  $\square$

**Lemma 4.14.** *The subspace of ergodic periodic measures in  $\mathcal{M}_G(\Omega^G)$  is dense in  $\mathcal{M}_G(\Omega^G)$ .*

*Proof.* Given two pmp actions  $G \curvearrowright (X, \lambda), G \curvearrowright (Y, \nu)$ , we say that the first action is *weakly contained* in the second action if for every finite sequence  $X_1, \dots, X_n \subset X$  of Borel sets, every finite sequence  $g_1, \dots, g_m \in G$  and every  $\epsilon > 0$  there exists a sequence  $Y_1, \dots, Y_n \in Y$  of Borel sets such that

$$|\lambda(X_j \cap g_i X_k) - \nu(Y_j \cap g_i Y_k)| < \epsilon \quad \forall 1 \leq i \leq m \text{ and } 1 \leq j, k \leq n.$$

This notion is due to Kechris [Ke12]. Because  $G$  is a free group, [Ke12, Theorem 3.1] implies that the canonical action of  $G$  on its profinite completion weakly contains every pmp action of  $G$ .

There is a slightly stronger characterization of weak containment. If  $G \curvearrowright (X, \lambda)$  is weakly contained in  $G \curvearrowright (Y, \nu)$  then for every finite sequence  $X_1, \dots, X_n \subset X$  of Borel sets, every finite set  $F \subset G$  and every  $\epsilon > 0$  there exists a sequence  $Y_1, \dots, Y_n \in Y$  of Borel sets such that for every function  $\phi : F \rightarrow \{1, \dots, n\}$ ,

$$\left| \lambda \left( \bigcap_{f \in F} f X_{\phi(f)} \right) - \nu \left( \bigcap_{f \in F} f Y_{\phi(f)} \right) \right| \leq \epsilon.$$

We leave this as an exercise to the reader. However, here is a hint: apply the original definition to the collection  $X_1, \dots, X_n, \{\bigcap_{f \in F} f X_{\phi(f)}\}_\phi$  and  $F^{-1}$ , using  $\epsilon/|F|$  in place of  $\epsilon$ .

For each  $\omega \in \Omega$ , let  $X_\omega := \{x \in \Omega^G : x(e) = \omega\}$  and  $\xi := \{X_\omega : \omega \in \Omega\}$  be the canonical partition of  $\Omega^G$ . Given a finite set  $F \subset G$ , we let  $\xi^F = \bigvee_{f \in F} f\xi$  be the common refinement. Given a function  $\phi : F \rightarrow \Omega$ , define

$$X_\phi = \bigcap_{f \in F} f X_{\phi(f)}.$$

So  $\xi^F = \{X_\phi : \phi : F \rightarrow \Omega\}$ .

Now let  $\lambda \in \mathcal{M}_G(\Omega^G)$ . Because the characteristic functions of sets in  $\xi^F$  (over all finite subsets  $F \subset G$ ) generate a dense sub-algebra of the space of continuous functions on  $\Omega^G$ , it suffices to show that for every finite set  $F \subset G$  and  $\epsilon > 0$  there exists an ergodic periodic measure  $\eta \in \mathcal{M}_G(\Omega^G)$  such that  $|\lambda(X) - \eta(X)| < \epsilon \quad \forall X \in \xi^F$ .

By Kechris' Theorem, if  $\overline{G}$  denotes the profinite completion of  $G$  then there exist sets  $\{G_\omega \subset \overline{G} : \omega \in \Omega\}$  such that

$$|\lambda(X_\phi) - \nu(G_\phi)| < \frac{\epsilon}{2|F|}$$

for every  $\phi : F \rightarrow \Omega$  where  $G_\phi = \bigcap_{f \in F} f G_{\phi(f)}$  and  $\nu$  is the Haar probability measure on  $\overline{G}$ .

By [Ke10, Lemma 10.2], we may assume without loss of generality that  $\{G_\omega : \omega \in \Omega\}$  is a partition of  $\overline{G}$ .

Let  $\{N_i\}_{i=1}^\infty$  be a decreasing sequence of finite-index normal subgroups of  $G$  such that  $\bigcap_{i=1}^\infty N_i = \{e\}$ . Moreover, we require that if  $K$  is any finite-index subgroup of  $G$  then  $N_i < K$  for all sufficiently large  $i$ .

Let  $\overline{N}_i$  be the closure of  $N_i$  in  $\overline{G}$  and  $\beta_i = \{g\overline{N}_i : g \in G\}$ . Then  $\{\beta_i\}$  is a decreasing sequence of finite partitions of  $\overline{G}$  and  $\bigvee_{i=1}^{\infty} \beta_i$  is the partition into points. So there is some  $n$  and a finite partition  $\alpha$  with  $\alpha \leq \beta_n$  such that  $\alpha = \{A_\omega : \omega \in \Omega\}$  and

$$\nu(A_\omega \Delta G_\omega) < \frac{\epsilon}{2|F|} \quad \forall \omega \in \Omega$$

where  $\Delta$  denotes symmetric difference. This implies  $|\nu(A_\phi) - \lambda(X_\phi)| < \epsilon$  for every  $\phi : F \rightarrow \Omega$  where  $A_\phi = \bigcap_{f \in F} f A_{\phi(f)}$ .

Define  $x : G \rightarrow \Omega$  by  $x(g) = \omega$  if and only if  $N_n g \subset A_\omega$ . Because  $\alpha \leq \beta_n$ , this is well-defined and  $x$  has a finite-index stabilizer. Let  $\eta$  be the uniform probability measure on the orbit of  $x \in \Omega^G$ . Then  $\eta \in \mathcal{M}_G(\Omega^G)$  is ergodic, periodic and  $\eta(X_\phi) = \nu(A_\phi)$  for every  $\phi : F \rightarrow \Omega$ . Thus  $|\lambda(X) - \eta(X)| < \epsilon \forall X \in \xi^F$  as required.  $\square$

As promised we can now prove that the  $\mathcal{M}_{\eta_n}^e(\widetilde{\text{Tree}_G})$  is pathwise connected. For this, recall that a convex closed metrizable subset  $\mathcal{K}$  of a locally convex linear space is a simplex if each point in  $\mathcal{K}$  is the barycenter of a unique probability measure supported on the subset  $\partial_e \mathcal{K}$  of extreme points of  $\mathcal{K}$ . In this case,  $\mathcal{K}$  is called a *Poulsen simplex* if  $\partial_e \mathcal{K}$  is dense in  $\mathcal{K}$ . It is known from [LOS78] that there is a unique Poulsen simplex up to affine isomorphism. Moreover, its set of extreme points is homeomorphic to  $l^2$ . The previous lemma immediately implies:

**Corollary 4.15.** *For each  $n \geq 1$ ,  $\mathcal{M}_{\eta_n}(\widetilde{\text{Tree}_G})$  is a Poulsen simplex. Therefore, the subspace of ergodic measures  $\mathcal{M}_{\eta_n}^e(\widetilde{\text{Tree}_G}) \subset \mathcal{M}_{\eta_n}(\widetilde{\text{Tree}_G})$  is homeomorphic to the Hilbert space  $l^2$ . In particular, it is pathwise connected.*

*Proof of Theorem 1.2.* By Corollary 4.7 and Lemma 4.13, for every  $n > 0$ , the set of all numbers  $t$  such that  $t = h_\mu(\nu_\lambda)$  for some ergodic periodic  $\lambda \in \mathcal{M}_{\eta_n}(\widetilde{\text{Tree}_G})$  is dense in  $[h_\mu(\eta_n), h_{\max}(\mu)]$ . By Lemma 4.12, for each such  $\lambda$  there exists a normal subgroup  $N \triangleleft G$  such that  $h_\mu(\nu_\lambda) = h_\mu(B_N, \nu_N)$ . By Lemma 4.8,  $\lim_{n \rightarrow \infty} h_\mu(\eta_n) = 0$ . This implies the theorem.  $\square$

## References

- [ABBGNRS11] M. Abert, N. Bergeron, I. Biringer, T. Gelander, N. Nikolov, J. Raimbault, I. Samet, *On the growth of Betti numbers of locally symmetric spaces*. arXiv:1104.55
- [AGV12] M. Abert, Y. Glasner, B. Virag, *Kesten's theorem for Invariant Random Subgroups*. arXiv:1201.3399
- [AL05] J. Aaronson, M. Lemańczyk, *Exactness of Rokhlin endomorphisms and weak mixing of Poisson boundaries*. Algebraic and topological dynamics, 77–87, Contemp. Math., 385, Amer. Math. Soc., Providence, RI, 2005.



- [BFLM11] J. Bourgain, A. Furman, E. Lindenstrauss, S. Mozes, *Stationary measures and equidistribution for orbits of nonabelian semigroups on the torus*. J. Amer. Math. Soc. 24 (2011), no. 1, 231–280.
- [BQ09] Y. Benoist, J.F. Quint, *Mesures stationnaires et fermés invariants des espaces homogènes*. C. R. Math. Acad. Sci. Paris 347 (2009), no. 1-2, 9–13,
- [BQ11a] Y. Benoist, J.F. Quint, *Mesures stationnaires et fermés invariants des espaces homogènes II*. C. R. Math. Acad. Sci. Paris 349 (2011), no. 5-6, 341–345
- [BQ11b] Y. Benoist, J.F. Quint, *Mesures stationnaires et fermés invariants des espaces homogènes*. Ann. of Math. (2) 174 (2011), no. 2, 1111–1162.
- [Bo03] L. Bowen. *Periodicity and Circle Packing in the Hyperbolic Plane*. Geom. Dedicata 102 (2003), 213–236.
- [Bo12] L. Bowen. *Invariant random subgroups of free groups*. preprint.
- [BS06] U. Bader, Y. Shalom, *Factor and normal subgroup theorems for lattices in products of groups*. Invent. Math. 163 (2006), 415–454
- [DS02] S. G. Dani, *On conjugacy classes of closed subgroups and stabilizers of Borel actions of Lie groups*. Ergodic Theory Dynam. Systems 22 (2002), no. 6, 1697–1714.
- [FG10] H. Furstenberg and E. Glasner. *Stationary dynamical systems*. Dynamical numbersinterplay between dynamical systems and number theory, 1–28, Contemp. Math., 532, Amer. Math. Soc., Providence, RI, (2010).
- [Fu80] H. Furstenberg. *Random walks on Lie Groups*, in: Wolf J. A., de Weild M.(Eds.), Harmonic Analysis and Representations of Semi-Simple Lie groups, D. Reidel, Dordrecht, 1980, pp. 467–489.
- [Fu77] H. Furstenberg. *Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions*. J. d’Analyse Math. 31, (1977), 204–256.
- [Fu72] H. Furstenberg. *Boundary theory and stochastic processes on homogeneous spaces, Harmonic analysis on homogeneous spaces* (Proc. Sympos. Pure Math., Vol. XXVI, Williams Coll., Williamstown, Mass., 1972), 193–229. Amer. Math. Soc., Providence, R.I., 1973.
- [Fu71] H. Furstenberg. *Random walks and discrete subgroups of Lie groups*. Advances in probability and related topics, Vol 1, Dekkers, 1971, pp. 1–63.
- [Fu63a] H. Furstenberg. *Noncommuting random products*. Trans. Amer. Math. Soc. 108 (1963) 377–428.
- [Fu63b] H. Furstenberg. *A Poisson formula for semi-simple Lie groups*. Ann. of Math. 77, (1963), 335–386.

- [GW97] E. Glasner, B. Weiss, Kazhdans property T and the geometry of the collection of invariant measures, *Geometry and Functional Analysis*, vol. 7 (1997), 917–935.
- [Gr11] R. Grigorchuk, *Some topics of dynamics of group actions on rooted trees.*, The Proceedings of the Steklov Institute of Math., v. 273 (2011), 1–118.
- [GS99] V. Golodets and S. D. Sinelshchikov, *On the conjugacy and isomorphism problems for stabilizers of Lie group actions.* *Ergodic Theory Dynam. Systems* 19 (1999), no. 2, 391–411.
- [KV83] V. A. Kaimanovich and A. M. Vershik. *Random walks on discrete groups: boundary and entropy.* *Ann. Probab.* 11 (1983), no. 3, 457–490.
- [Ke10] A. Kechris, *Global aspects of ergodic group actions.* Mathematical Surveys and Monographs, 160. American Mathematical Society, Providence, RI, 2010. xii+237
- [Ke12] A. S. Kechris. *Weak containment in the space of actions of a free group*, *Israel J. Math.*, to appear.
- [LOS78] J. Lindenstrauss, G.H. Olsen, Y. Sternfeld, The Poulsen simplex, *Ann. Inst. Fourier (Grenoble)* 28 (1978), 91–114.
- [Ne03] A. Nevo. *The spectral theory of amenable actions and invariants of discrete groups.* *Geom. Dedicata* 100 (2003), 187–218.
- [NZ02a] A. Nevo and R. Zimmer. *A structure theorem for actions of semisimple Lie groups.* *Ann. of Math. (2)* 156 (2002), no. 2, 565–594.
- [NZ02b] A. Nevo and R. Zimmer. *A generalization of the intermediate factors theorem.* *J. Anal. Math.* 86 (2002), 93–104.
- [NZ00] A. Nevo and R. Zimmer. *Rigidity of Furstenberg entropy for semisimple Lie group actions.* *Ann. Sci. École Norm. Sup. (4)* 33 (2000), no. 3, 321–343.
- [NZ99] A. Nevo and R. Zimmer. *Homogenous projective factors for actions of semi-simple Lie groups.* *Invent. Math.* 138 (1999), no. 2, 229–252.
- [Sa11] D. Savchuk, *Schreier Graphs of Actions of Thompsons Group F on the Unit Interval and on the Cantor Set*, arXiv: 1105.4017.
- [SZ94] G. Stuck and R. J. Zimmer, *Stabilizers for ergodic actions of higher rank semisimple groups*, *Ann. of Math. (2)* 139 (1994), no. 3, 723–747.
- [Ve10] A. Vershik, *Nonfree Actions of Countable Groups and their Characters*, arXiv:1012.4604
- [Ve11] A. Vershik, *Totally nonfree actions and infinite symmetric group.* arXiv:1109.3413
- [Vo12] Y. Vorobets, *Notes on the Schreier graphs of the Grigorchuk group.* *Contemporary Mathematics* 567 (2012), 221–248.